GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 10, 1999, 06:14:03 ; Search time 9.95 Seconds
(without alignments)
488.658 Million cell updates/sec

Title: US-09-208-619-1
Perfect score: 911
Sequence: 1 MEEYAREPCPWRIVDDCGGA.....QLPPKDGTPAPGYPSYQQYH 172

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

I[redacted]e : SwissProt_37:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	704.5	77.3	171	1	IM17_RAT	O35092 rattus norv
2	703.5	77.2	171	1	IM17_HUMAN	Q99595 homo sapien
3	377.5	41.4	158	1	IM17_YEAST	P39515 saccharomyc
4	343.5	37.7	164	1	IM17_SCHPO	P87130 schizosacch
5	117	12.8	175	1	IM22_SCHPO	P87146 schizosacch
6	115.5	12.7	207	1	IM22_YEAST	Q12328 saccharomyc
7	104	11.4	222	1	IM23_YEAST	P32897 saccharomyc
8	89.5	9.8	200	1	NUJM_NEUCR	P25710 neurospora
9	86	9.4	1262	1	CYA5_RAT	Q04400 rattus norv
10	83	9.1	574	1	KPYR_HUMAN	P30613 homo sapien
11	83	9.1	541	1	PTN5_MOUSE	P54830 mus musculu
12	80.5	8.8	417	1	GLYA_BORBU	O51547 borrelia bu
13	80.5	8.8	256	1	Y058_MYCTU	Q10887 mycobacteri
14	80	8.8	1123	1	DC11_DROME	P18169 drosophila
15	79	8.7	674	1	PTGA_BRELA	Q45298 brevibacter
16	78.5	8.6	1264	1	CYA5_RABIT	P40144 oryctolagus
17	77	8.5	519	1	KPYR_CANFA	Q29536 canis famil
18	76.5	8.4	347	1	TCR_STRRM	P14551 streptomyce
19	76	8.3	495	1	ACCD_MYCTU	Q10561 mycobacteri
20	75	8.2	324	1	YAGC_SCHPO	Q09875 schizosacch
21	74.5	8.2	356	1	HRPN_ERWCA	Q47279 erwinia car
22	74.5	8.2	456	1	YAJR_ECOLI	P77726 escherichia
23	73.5	8.1	466	1	HYIN_AGRRA	Q09102 agrobacteri
24	73.5	8.1	416	1	YBDA_ECOLI	P24077 escherichia
25	73	8.0	537	1	PTN5_HUMAN	P54829 homo sapien
26	73	8.0	503	1	PUR8_STRLP	P42670 streptomyce
27	72.5	8.0	976	1	EPA2_HUMAN	P29317 homo sapien
28	72.5	8.0	854	1	KDPD_RATRA	O34971 rathayibact
29	72.5	8.0	538	1	TCMA_STRGA	P39886 streptomyce
30	72.5	8.0	854	1	VGLB_RHCM6	P89053 rhesus cyto
31	72	7.9	247	1	SUMT_PSEFL	P37725 pseudomonas
32	71.5	7.8	574	1	KPYR_MOUSE	P53657 mus musculu
33	71.5	7.8	574	1	KPYR_RAT	P12928 rattus norv
34	71.5	7.8	288	1	M6B_MOUSE	P35803 mus musculu
35	71	7.8	750	1	ELS_CHICK	P07916 gallus gall
36	70.5	7.7	661	1	HS7C_CAEEL	P27420 caenorhabdi
37	70.5	7.7	631	1	YF14_HAEIN	P44236 haemophilus
38	70	7.7	233	1	MTRD_METTM	P80183 methanobact
39	70	7.7	296	1	SAPC_ECOLI	Q47624 escherichia
40	69.5	7.6	261	1	AAC4_SALSP	P08988 salmonella
41	69.5	7.6	314	1	YM78_YEAST	Q04013 saccharomyc
42	69	7.6	324	1	PMT_YEAST	P32332 saccharomyc
43	69	7.6	284	1	THT3_MYCTU	Q59570 mycobacteri
44	69	7.6	406	1	TYRQ_HAEIN	P44747 haemophilus
45	69	7.6	375	1	XYLH_HAEIN	P45045 haemophilus

ALIGNMENTS

RESULT 1
IM17_RAT
ID IM17_RAT STANDARD; PRT; 171 AA.
AC O35092;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM17 HOMOLOG.
GN TIM17.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA ISHIHARA N., MIHARA K.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -!- FUNCTION: COMPONENT OF THE PREPROTEIN IMPORT MACHINERY OF THE
CC MITOCHONDRIAL INNER MEMBRANE. INTEGRAL PART OF A PUTATIVE PROTEIN-
CC CONDUCTING CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE (BY
CC SIMILARITY).
CC -!- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT
CC LEAST 3 DIFFERENT PROTEINS (TIM17, TIM23, TIM44) (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC INNER MEMBRANE (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC ---

CC ---
DR EMBL; AB006450; D1022674; -.
KW PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT TRANSMEM 17 37 POTENTIAL.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 113 133 POTENTIAL.
SQ SEQUENCE 171 AA; 18037 MW; 6E765C35 CRC32;

Query Match 77.3%; Score 704.5; DB 1; Length 171;
Best Local Similarity 75.4%; Pred. No. 3.4e-57;
Matches 129; Conservative 17; Mismatches 24; Indels 1; Gaps 1;

```
Qy       1 MEEYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAP 60
           |||||||||||||||||||||||| ||||:||| |||||:|||: ||||||  |:: |||
Db       1 MEEYAREPCPWRIVDDCGGAFTMGTIGGGIFQAFKGFRNSPVGVNHRLRGSLTAIKTRAP 60

Qy      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
             |:|||||||||| |||||:|::|||||||||||||||||:||||:||:|||||| |||
Db      61 QLGGSFAVWGGLFSTIDCGMVQIRGKEDPWNSITSGALTGAILAARNGPVAMVGSAAMGG 120

Qy     121 ILLALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSYQQY 171
           |||||||| |||||: : || | | | || ||||     |:  :  |:||
Db     121 ILLALIEGAGILLTRFASAQFPNGPQFAEDHSQLPSSQ-LPSSPFGDYRQY 170
```

RESULT 2
IM17_HUMAN
ID IM17_HUMAN STANDARD; PRT; 171 AA.
AC Q99595;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

```
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM17 HOMOLOG.
GN   TIM17.
OS   HOMO SAPIENS (HUMAN).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC   PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN   [1]
RP   SEQUENCE FROM N.A.
RA   BOEMER U., RASSOW J., PFANNER N., MEIJER M., MAARSE A.C.;
RL   SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: COMPONENT OF THE PREPROTEIN IMPORT MACHINERY OF THE
CC       MITOCHONDRIAL INNER MEMBRANE. INTEGRAL PART OF A PUTATIVE PROTEIN-
CC       CONDUCTING CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT
CC       LEAST 3 DIFFERENT PROTEINS (TIM17, TIM23, TIM44) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X97544; E243529; -.
KW   PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT   TRANSMEM      17     37       POTENTIAL.
FT   TRANSMEM      63     83       POTENTIAL.
FT   TRANSMEM     113    133       POTENTIAL.
SQ   SEQUENCE   171 AA;  18023 MW;  6408C99C CRC32;
```

```
  Query Match           77.2%;  Score 703.5;  DB 1;  Length 171;
  Best Local Similarity  75.4%;  Pred. No. 4.2e-57;
  Matches  129;  Conservative   17;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 MEEYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAP 60
           |||||||||||||||||||||||| ||||:|||||||||:|||: |||||  |:: |||
Db       1 MEEYAREPCPWRIVDDCGGAFTMGTIGGGIFQAIKGFRNSPVGVNHRLRGSLTAIKTRAP 60

Qy      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
             |:|||||||||||  ||| :|::|||||||||||||||||:||||:||:|||||| |||
Db      61 QLGGGSFAVWGGLFSMIDCSMVQVRGKEDPWNSITSGALTGAILAARNGPVAMVGSAAMGG 120

Qy     121 ILLALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSYQQY 171
           |||||||| ||||||: : || | | | ||||||||     |:  :  |:||
Db     121 ILLALIEGAGILLTRFASAQFPNGPQFAEDPSQLPSTQ-LPSSPFGDYRQY 170
```

```
RESULT   3
IM17_YEAST
ID   IM17_YEAST     STANDARD;      PRT;   158 AA.
AC   P39515; Q02310;
DT   01-FEB-1995 (REL. 31, CREATED)
DT   01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM17
DE   (MITOCHONDRIAL PROTEIN IMPORT PROTEIN 2) (MITOCHONDRIAL INNER MEMBRANE
DE   PROTEIN MIM17).
GN   TIM17 OR MPI2 OR MIM17 OR SMS1 OR YJL143W OR J0648.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC   SACCHAROMYCETACEAE; SACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 77162;
RX   MEDLINE; 94326923.
RA   MAARSE A.C., BLOM J., KEIL P., PFANNER N., MEIJER M.;
RT   "Identification of the essential yeast protein MIM17, an integral
RT   mitochondrial inner membrane protein involved in protein import.";
RL   FEBS LETT. 349:215-221(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 95003211.
RA   RYAN K.R., MENOLD M.M., GARRETT S., JENSEN R.E.;
RT   "SMS1, a high-copy suppressor of the yeast mas6 mutant, encodes an
RT   essential inner membrane protein required for mitochondrial protein
RT   import.";
RL   MOL. BIOL. CELL 5:529-538(1994).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RX   MEDLINE; 96408771.
RA   KATSOULOU C., TZERMIA M., TAVERNARAKIS N., ALEXANDRAKI D.;
RT   "Sequence analysis of a 40.7 kb segment from the left arm of yeast
RT   chromosome X reveals 14 known genes and 13 new open reading frames
RT   including homologues of genes clustered on the right arm of
RT   chromosome XI.";
RL   YEAST 12:787-797(1996).
RN   [4]
RP   CHARACTERIZATION, AND TOPOLOGY.
RX   MEDLINE; 94326924.
RA   KUEBRICH M., KEIL P., RASSOW J., DEKKER P.J.T., BLOM J., MEIJER M.,
RA   PFANNER N.;
RT   "The polytopic mitochondrial inner membrane proteins MIM17 and MIM23
RT   operate at the same preprotein import site.";
RL   FEBS LETT. 349:222-228(1994).
CC   -!- FUNCTION: COMPONENT OF THE PREPROTEIN IMPORT MACHINERY OF
CC       THE MITOCHONDRIAL INNER MEMBRANE. INTEGRAL PART OF A PUTATIVE
CC       PROTEIN-CONDUCTING CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT
CC       LEAST 3 DIFFERENT PROTEINS (TIM17, TIM23, TIM44).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X77796; G557267; -.
DR   EMBL; S74018; G693752; -.
DR   EMBL; X87371; G854554; -.
DR   EMBL; Z49418; G1015553; -.
DR   SGD; L0001139; TIM17.
KW   PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT   TRANSMEM      12     32       POTENTIAL.
FT   TRANSMEM      59     79       POTENTIAL.
FT   TRANSMEM      88    108       POTENTIAL.
FT   TRANSMEM     113    133       POTENTIAL.
FT   CONFLICT      33     33       H -> T (IN REF. 2).
SQ   SEQUENCE   158 AA;  16584 MW;  8BDDCBFA CRC32;
```

```
  Query Match           41.4%;  Score 377.5;  DB 1;  Length 158;
  Best Local Similarity  48.7%;  Pred. No. 1.3e-27;
  Matches   75;  Conservative   26;  Mismatches   50;  Indels    3;  Gaps    2;

Qy       3 EYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQI 62
           :::|:|||  |::| |||| || ||| |:  ||||||:|:|   |  |: :|:: ||| :
Db       4 DHSRDPCPIVILNDFGGAFAMGAIGGVVWHGIKGFRNSPLG--ERGSGAMSAIKARAPVL 61

Qy      63 GGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGGIL 122
           ||:| |||||| | || :  :| :|||||:| :|  ||  || | |      |::    |
Db      62 GGNFGVWGGLFSTFDCAVKAVRKREDPWNAIIAGFFTGGALAVRGGWRHTRNSSITCACL 121

Qy     123 LALIEGVGILLTRYTAQQFR-NAPPFLEDPSQLP 155
           | :||||::  || | | :  |||  | ||  |
```

```
Db    122 LGVIEGVGLMFQRYAAWQAKPMAPPLPEAPSSQP 155


RESULT   4
IM17_SCHPO
ID   IM17_SCHPO     STANDARD;      PRT;   164 AA.
AC   P87130;
DT   15-JUL-1998 (REL. 36, CREATED)
DT   15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM17 HOMOLOG
DE   (MITOCHONDRIAL PROTEIN IMPORT PROTEIN 2) (MITOCHONDRIAL INNER MEMBRANE
DE   PROTEIN MIM17).
GN   SPAC3A12.16C.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC   SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC   SCHIZOSACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   BADCOCK K., CHURCHER C.M., WOOD V., BARRELL B.G., RAJANDREAM M.A.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: COMPONENT OF THE PREPROTEIN IMPORT MACHINERY OF THE
CC       MITOCHONDRIAL INNER MEMBRANE. INTEGRAL PART OF A PUTATIVE PROTEIN-
CC       CONDUCTING CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT
CC       LEAST 3 DIFFERENT PROTEINS (TIM17, TIM23, TIM44) (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z95395; E316107; -.
KW   PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT   TRANSMEM     13     33       POTENTIAL.
FT   TRANSMEM     60     80       POTENTIAL.
FT   TRANSMEM     89    109       POTENTIAL.
FT   TRANSMEM    114    134       POTENTIAL.
SQ   SEQUENCE   164 AA;  16901 MW;  044D336B CRC32;


  Query Match             37.7%;  Score 343.5;  DB 1;  Length 164;
  Best Local Similarity   43.6%;  Pred. No. 1.6e-24;
  Matches   72;  Conservative   26;  Mismatches   56;  Indels   11;  Gaps    3;

Qy       3 EYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQI 62
           :: |:|||: |::| | ||:|| ||| :: :|||:||:| |   |:   | | : ||| :
Db       5 DHTRDPCPYVILNDFGAAFSMGTIGGAIWHSIKGWRNSPPG-EKRISAIA-AAKTRAPVL 62

Qy      63 GGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGGIL 122
           ||:| |||||| | || :  :| ||||||:| :|  ||  || | |  |    |:    :
Db      63 GGNFGVWGGLFSTFDCAVKGVRRKEDPWNAIIAGFFTGGALAVRGGWRATRNGAIGCACI 122

Qy     123 LALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPS 167
           ||: ||:|| | |  |:  |   | : |          || |  |
Db     123 LAVFEGLGIALGRMNAEYNRPVAPVIPD---------APASGSTS 158


RESULT   5
IM22_SCHPO
ID   IM22_SCHPO     STANDARD;      PRT;   175 AA.
AC   P87146;
DT   15-JUL-1998 (REL. 36, CREATED)
DT   15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM22 HOMOLOG.
GN   SPBC25H2.04C.
OS   SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; ARCHIASCOMYCETES;
OC   SCHIZOSACCHAROMYCETALES; SCHIZOSACCHAROMYCETACEAE;
OC   SCHIZOSACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RA   DURSO G., LYE G., BOWMAN S., CHURCH C., WOOD V., BARRELL B.G.,
RA   RAJANDREAM M.A., CONNOR R.E.;
RL   SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF MITOCHONDRIAL INNER MEMBRANE
CC       PROTEIN IMPORT PATHWAY. INVOLVED IN THE IMPORT OF PROTEINS
CC       TARGETED TO THE INNER MEMBRANE OF THE MITOCHONDRION SUCH AS
CC       MITOCHONDRIAL CARRIER FAMILY MEMBERS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z95397; E315883; -.
KW   PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT   TRANSMEM     51     71       POTENTIAL.
FT   TRANSMEM    126    146       POTENTIAL.
FT   TRANSMEM    153    173       POTENTIAL.
SQ   SEQUENCE   175 AA;  18641 MW;  BE82B517 CRC32;


  Query Match             12.8%;  Score 117;  DB 1;  Length 175;
  Best Local Similarity   26.7%;  Pred. No. 0.00053;
  Matches   39;  Conservative   26;  Mismatches   49;  Indels   32;  Gaps    6;

Qy      12 RIVDDCGGAFTM-GVIG---GGVFQAIKGFRNAPVGIRH-----RLRGSANAVRIRAPQI 62
           || : |    :| ||:|   ||:|    |   : : ::|       :     :||:|  :
Db      41 RISESCVFKSSMAGVLGFGLGGIF----GMFISSLDLQHIDPKIYEKPFREQIRIQARDM 96

Qy      63 GG-------SFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGS 115
           |        :| : | ::   :| :   | | | :|:| :|  ||  || |||| |:|
Db      97 GSRSFSTAKNFGLLGLIYSGSECCIEAFRAKTDIYNAIAAGVFTGGALAVRSGPKAIV-- 155

Qy     116 AMMGGILLALIEGVGILLTRYTAQQF 141
                 : | |  |  |  :::
Db     155 ----------LGGAGFGLFSYGIEKY 170


RESULT   6
IM22_YEAST
ID   IM22_YEAST     STANDARD;      PRT;   207 AA.
AC   Q12328;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM22.
GN   TIM22 OR YDL217C OR D0884.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC   SACCHAROMYCETACEAE; SACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C / FY1679;
RA   BAHR A., MOELLER-RIEKER S., HANKELN T., KRAEMER C., SCHMIDT E.R.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE FROM N.A.
```

```
RA   RASMUSSEN S.W.;
RL   SUBMITTED (JUL-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [3]
RP   CHARACTERIZATION.
RX   MEDLINE; 97113414.
RA   SIRRENBERG C., BAUER M.D., GUIARD B., NEUPERT W., BRUNNER M.;
RT   "Import of carrier proteins into the mitochondrial inner membrane
RT   mediated by Tim22.";
RL   NATURE 384:582-585(1996).
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF MITOCHONDRIAL INNER MEMBRANE
CC       PROTEIN IMPORT PATHWAY. INVOLVED IN THE IMPORT OF PROTEINS
CC       TARGETED TO THE INNER MEMBRANE OF THE MITOCHONDRION SUCH AS
CC       MITOCHONDRIAL CARRIER FAMILY MEMBERS.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE TIM17/TIM22 FAMILY.
CC   ---------------------------------------------------------------------
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X99000; E253394; -.
DR   EMBL; Z74265; E253267; -.
DR   SGD; L0004157; TIM22.
KW   PROTEIN TRANSPORT; TRANSMEMBRANE; MITOCHONDRION; INNER MEMBRANE.
FT   TRANSMEM     47     67       POTENTIAL.
FT   TRANSMEM    152    172       POTENTIAL.
FT   TRANSMEM    173    193       POTENTIAL.
SQ   SEQUENCE   207 AA;  21864 MW;  AA8B8201 CRC32;
```

```
  Query Match             12.7%;  Score 115.5;  DB 1;  Length 207;
  Best Local Similarity   30.0%;  Pred. No. 0.00086;
  Matches   30;  Conservative   23;  Mismatches   40;  Indels    7;  Gaps    3;

Qy      22 TMGVIG-GGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQIGGSFAVWGGLFXTIDCGL 80
           | | || ||: : ::   : |   | :::     :  ::     :|   | ::  ::| :
Db      86 TAGNIGVGGISRTVQQISDLP--FRQQMKLQFTDMGKKSYSSAKNFGYIGMIYAGVECVI 143

Qy      81 VRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
             || | | :| :|:|  ||| || ::|| |    |:|||
Db     144 ESLRAKNDIYNGVTAGFFTGAGLAYKAGPQA----ALMGG 179
```

```
RESULT   7
YEAST
ID   IM23_YEAST     STANDARD;      PRT;   222 AA.
AC   P32897;
DT   01-OCT-1993 (REL. 27, CREATED)
DT   01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   MITOCHONDRIAL IMPORT INNER MEMBRANE TRANSLOCASE SUBUNIT TIM23
DE   (MITOCHONDRIAL PROTEIN IMPORT PROTEIN 3) (MITOCHONDRIAL PROTEIN IMPORT
DE   PROTEIN MAS6) (MEMBRANE IMPORT MACHINERY PROTEIN MIM23).
GN   TIM23 OR MAS6 OR MPI3 OR MIM23 OR YNR017W OR N3180.
OS   SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;
OC   SACCHAROMYCETACEAE; SACCHAROMYCES.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93359499.
RA   EMTAGE J.L.T., JENSEN R.E.;
RT   "MAS6 encodes an essential inner membrane component of the yeast
RT   mitochondrial protein import pathway.";
RL   J. CELL BIOL. 122:1003-1012(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 93380573.
RA   DEKKER P.J.T., KEIL P., RASSOW J., MAARSE A.C., PFANNER N.,
RA   MEIJER M.;
RT   "Identification of MIM23, a putative component of the protein import
RT   machinery of the mitochondrial inner membrane.";
RL   FEBS LETT. 330:66-70(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   POHL T.M.;
RL   SUBMITTED (MAY-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [4]
RP   CHARACTERIZATION, AND TOPOLOGY.
RX   MEDLINE; 94326924.
RA   KUEBRICH M., KEIL P., RASSOW J., DEKKER P.J.T., BLOM J., MEIJER M.,
RA   PFANNER N.;
RT   "The polytopic mitochondrial inner membrane proteins MIM17 and MIM23
RT   operate at the same preprotein import site.";
RL   FEBS LETT. 349:222-228(1994).
CC   -!- FUNCTION: ESSENTIAL COMPONENT OF MITOCHONDRIAL MATRIX PROTEIN
CC       IMPORT PATHWAY. PROBABLY INVOLVED IN TRANSLOCATION ACROSS THE
CC       INNER MEMBRANE. INTEGRAL PART OF A PUTATIVE PROTEIN-CONDUCTING
CC       CHANNEL OF THE MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SUBUNIT: FORMS PART OF THE RECEPTOR COMPLEX THAT CONSISTS OF AT
CC       LEAST 3 DIFFERENT PROTEINS (TIM17, TIM23, TIM44).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
CC       INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE TIM23 FAMILY.
CC   ---------------------------------------------------------------------
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X71633; G297118; -.
DR   EMBL; X74161; G402785; -.
DR   EMBL; Z71632; E239788; -.
DR   PIR; S33201; S33201.
DR   PIR; S36139; S36139.
DR   PIR; A40705; A40705.
DR   SGD; L0001028; MAS6.
KW   MITOCHONDRION; INNER MEMBRANE; TRANSPORT; PROTEIN TRANSPORT;
KW   TRANSMEMBRANE; TRANSLOCATION.
FT   DOMAIN        1     96       INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM     97    130       POTENTIAL.
FT   DOMAIN      131    144       MATRIX (POTENTIAL).
FT   TRANSMEM    145    166       POTENTIAL.
FT   DOMAIN      167    174       INTERMEMBRANE (POTENTIAL).
FT   TRANSMEM    175    189       POTENTIAL.
FT   DOMAIN      190    196       MATRIX (POTENTIAL).
FT   TRANSMEM    197    215       POTENTIAL.
FT   DOMAIN      214    222       INTERMEMBRANE (POTENTIAL).
SQ   SEQUENCE   222 AA;  23243 MW;  320C2078 CRC32;
```

```
  Query Match             11.4%;  Score 104;  DB 1;  Length 222;
  Best Local Similarity   29.9%;  Pred. No. 0.01;
  Matches   32;  Conservative   21;  Mismatches   46;  Indels    8;  Gaps    3;

Qy      18 GGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSA--NAVRIRAPQIGGSFAVWGGLFXT 75
           |  : :|:  ||    ::| :| |     :|: :   | :  | | :| :  :     :
Db     102 GAVYLLGLGIGGFSGMMQGLQNIPPNSPGKLQLNTVLNHITKRGPFLGNNAGILALSYNI 161

Qy      76 IDCGLVRLRGKEDPWNSITSGALTGAVLAARSG--PL----AMVGSA 116
           |:  :  |||| |   || :||||||:  :  |  |:    ||| :|
Db     162 INSTIDALRGKHDTAGSIGAGALTGALFKSSKGLKPMGYSSAMVAAA 208
```

```
RESULT   8
NUJM_NEUCR
ID   NUJM_NEUCR     STANDARD;      PRT;   200 AA.
AC   P25710;
DT   01-MAY-1992 (REL. 22, CREATED)
```

```
DT   01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
DT   01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE   NADH-UBIQUINONE OXIDOREDUCTASE 21.3 KD SUBUNIT (EC 1.6.5.3)
DE   (EC 1.6.99.3).
OS   NEUROSPORA CRASSA.
OC   EUKARYOTA; FUNGI; ASCOMYCOTA; EUASCOMYCETES; PYRENOMYCETES;
OC   SORDARIALES; SORDARIACEAE; NEUROSPORA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ST. LAWRENCE 74 / SL 74 / ORS 6A;
RX   MEDLINE; 91159490.
RA   NEHLS U., HEMMER S., ROEHLEN D.A., VAN DER PAS J.C., PREIS D.,
RA   SACKMANN U., WEISS H.;
RT   "cDNA and genomic DNA sequence of the 21.3 kDa subunit of
RT   NADH:ubiquinone reductase (complex I) from Neurospora crassa.";
RL   BIOCHIM. BIOPHYS. ACTA 1088:325-326(1991).
CC   -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC       CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC       TO BE UBIQUINONE.
CC   -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.
CC   -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.
CC   -!- SUBCELLULAR LOCATION: MATRIX AND CYTOPLASMIC SIDE OF THE
CC       MITOCHONDRIAL INNER MEMBRANE.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X56612; G3050; -.
DR   PIR; S14277; S14277.
KW   OXIDOREDUCTASE; NAD; UBIQUINONE; MITOCHONDRION.
SQ   SEQUENCE   200 AA;  21349 MW;  A0B12273 CRC32;


  Query Match             9.8%;  Score 89.5;  DB 1;  Length 200;
  Best Local Similarity   27.1%;  Pred. No. 0.19;
  Matches   32;  Conservative   14;  Mismatches   57;  Indels   15;  Gaps    3;

Qy      63 GG---SFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARS--GPLAMVGSAM 117
           ||   || : ||::          || ||| ||      | |:|:  ||   |: :   ||
Db      55 GGTIISFTLAGGIYRFAQQASANLREKEDGWNHAIGAFLGGSVMGLRSLRFPVILGFGAM 114

Qy     118 MGGILLALIEGVGIL----------LTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGY 165
           | :: |     |:             |  | :     |  |  :::    |   |||
Db     115 AGSVVGAFAFSGGLTGWGRDPNVDEFERKEAMRLNRRRPVEETLAEVGEGRGIYPPGY 172


RESULT   9
CYA5_RAT
ID   CYA5_RAT       STANDARD;      PRT;  1262 AA.
AC   Q04400;
DT   01-OCT-1993 (REL. 27, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   ADENYLATE CYCLASE, TYPE V (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE)
DE   (ADENYLYL CYCLASE).
GN   ADCY5.
OS   RATTUS NORVEGICUS (RAT).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC   RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER, KIDNEY, AND HEART;
RA   PREMONT R.T.;
RL   SUBMITTED (DEC-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
RN   [2]
RP   SEQUENCE OF 164-1262 FROM N.A.
RC   TISSUE=LIVER;
RX   MEDLINE; 93028552.
RA   PREMONT R.T., CHEN J., MA H.-W., PONNAPALLI M., IYENGAR R.;
RT   "Two members of a widely expressed subfamily of hormone-stimulated
RT   adenylyl cyclases.";
RL   PROC. NATL. ACAD. SCI. U.S.A. 89:9809-9813(1992).
CC   -!- FUNCTION: THIS A MEMBRANE-BOUND, CALMODULIN-INSENSITIVE ADENYLYL
CC       CYCLASE.
CC   -!- CATALYTIC ACTIVITY: ATP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.
CC   -!- ENZYME REGULATION: INSENSIBLE TO CA(2+)/CALMODULIN. STIMULATED BY
CC       THE G PROTEIN BETA & GAMMA SUBUNIT COMPLEX.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. CONTAINS TWO
CC       DOMAINS WITH 6 POTENTIAL TRANSMEMBRANE REGIONS EACH.
CC   -!- SIMILARITY: THE TWO CYTOPLASMIC DOMAINS ARE SIMILAR TO ONE
CC       ANOTHER AND TO THE CATALYTIC DOMAINS OF GUANYLATE CYCLASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M96159; G1758332; -.
DR   PROSITE; PS00452; GUANYLATE_CYCLASES; 2.
DR   PFAM; PF00211; guanylate_cyc; 2.
DR   HSSP; P19754; 1AWK.
KW   LYASE; CAMP SYNTHESIS; TRANSMEMBRANE; GLYCOPROTEIN; DUPLICATION.
FT   DOMAIN        1    242       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    243    263       POTENTIAL.
FT   TRANSMEM    269    289       POTENTIAL.
FT   TRANSMEM    300    320       POTENTIAL.
FT   TRANSMEM    326    346       POTENTIAL.
FT   TRANSMEM    351    371       POTENTIAL.
FT   TRANSMEM    375    395       POTENTIAL.
FT   DOMAIN      396    763       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM    764    784       POTENTIAL.
FT   TRANSMEM    790    810       POTENTIAL.
FT   TRANSMEM    837    857       POTENTIAL.
FT   DOMAIN      858    910       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    911    931       POTENTIAL.
FT   TRANSMEM    936    956       POTENTIAL.
FT   TRANSMEM    985   1005       POTENTIAL.
FT   DOMAIN     1006   1262       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   1262 AA;  139178 MW;  3F17897B CRC32;


  Query Match             9.4%;  Score 86;  DB 1;  Length 1262;
  Best Local Similarity   26.2%;  Pred. No. 2.7;
  Matches   42;  Conservative   15;  Mismatches   43;  Indels   60;  Gaps    7;

Qy      11 WRIVDDCGGAFTMG---VIGGGVFQAIKGFRNAPV----------GIRHRLRGSANAVR 56
           ||  ||       |   ::||  |     ||:             | | : ||:|
Db      73 WRGGDDDEDPPLSGDDPLVGGFGF----SFRSKSAWQERGGDDGGRGSRRQRRGAAGGGS 128

Qy      57 IRAPQIGGS------FAVWGGLF---XTIDCGLVRLRGK-----------------E 87
           |||  |||       |  ||      ::: ||   |||                 |
Db     129 TRAPPAGGSGSSAAAAAAAGGTEVRPRSVEVGLEERRGKGRAAEELEPGTGTVEDGDGSE 188

Qy      88 DPWNSITSGALTGAVLAARSGPLAMVGSAMMGGILLALIE 127
           |  :|: ||: || ||:          :|   |||::
Db     189 DGGSSVASGSGTGTVLS----------LGACCLALLQ 215


RESULT  10
KPYR_HUMAN
ID   KPYR_HUMAN     STANDARD;      PRT;   574 AA.
AC   P30613; P11973;
DT   01-APR-1993 (REL. 25, CREATED)
DT   01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT   15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE   PYRUVATE KINASE, ISOZYMES R/L (EC 2.7.1.40).
GN   PKLR OR PKL OR PK1.
OS   HOMO SAPIENS (HUMAN).
```

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A., AND VARIANT TOKYO.
RX MEDLINE; 91376115.
RA KANNO H., FUJII H., HIRONO A., MIWA S.;
RT "cDNA cloning of human R-type pyruvate kinase and identification of a
RT single amino acid substitution (Thr384-->Met) affecting enzymatic
RT stability in a pyruvate kinase variant (PK Tokyo) associated with
RT hereditary hemolytic anemia.";
RL PROC. NATL. ACAD. SCI. U.S.A. 88:8218-8221(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 88158079.
RA TANI K., FUJII H., NAGATA S., MIWA S.;
RT "Human liver type pyruvate kinase: complete amino acid sequence and
RT the expression in mammalian cells.";
RL PROC. NATL. ACAD. SCI. U.S.A. 85:1792-1795(1988).
RN [3]
RP SEQUENCE OF 470-574 FROM N.A.
RX MEDLINE; 87184521.
RA TANI K., FUJII H., TSUTSUMI H., SUKEGAWA J., TOYOSHIMA K.,
RA YOSHIDA M.C., NOGUCHI T., TANAKA T., MIWA S.;
RT "Human liver type pyruvate kinase: cDNA cloning and chromosomal
RT assignment.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 143:431-438(1987).
RN [4]
RP VARIANTS LINZ AND BEIRUT.
RX MEDLINE; 91208396.
RA NEUBAUER B., LAKOMEK M., WINKLER H., PARKE M., HOFFERBERT S.,
RA SCHROTER W.;
RT "Point mutations in the L-type pyruvate kinase gene of two children
RT with hemolytic anemia caused by pyruvate kinase deficiency.";
RL BLOOD 77:1871-1875(1991).
RN [5]
RP VARIANT FUKUSHIMA/MAEBASHI.
RX MEDLINE; 92163106.
RA KANNO H., FUJII H., HIRONO A., OMINE M., MIWA S.;
RT "Identical point mutations of the R-type pyruvate kinase (PK) cDNA
RT found in unrelated PK variants associated with hereditary hemolytic
RT anemia.";
RL BLOOD 79:1347-1350(1992).
RN [6]
RP SEQUENCE OF 365-431 FROM N.A., AND VARIANT OSAKA.
RX MEDLINE; 93236593.
RA KANNO H., FUJII H., TSUJINO G., MIWA S.;
RT "Molecular basis of impaired pyruvate kinase isozyme conversion in
RT erythroid cells: a single amino acid substitution near the active
RT site and decreased mRNA content of the R-type PK.";
RL BIOCHEM. BIOPHYS. RES. COMMUN. 192:46-52(1993).
RN [7]
RP REVIEW ON VARIANTS.
RX MEDLINE; 96263732.
RA BEUTLER E., BARONCIANI L.;
RT "Mutations in pyruvate kinase.";
RL HUM. MUTAT. 7:1-6(1995).
RN [8]
RP VARIANTS SAPPORO GLN-426.
RX MEDLINE; 93244440.
RA KANNO H., FUJII H., MIWA S.;
RT "Low substrate affinity of pyruvate kinase variant (PK Sapporo)
RT caused by a single amino acid substitution (426 Arg-->Gln) associated
RT with hereditary hemolytic anemia.";
RL BLOOD 81:2439-2441(1993).
RN [9]
RP VARIANTS D-134; P-155; H-359; W-486; V-495 AND Q-510.
RX MEDLINE; 93248282.
RA BARONCIANI L., BEUTLER E.;
RT "Analysis of pyruvate kinase-deficiency mutations that produce
RT nonspherocytic hemolytic anemia.";
RL PROC. NATL. ACAD. SCI. U.S.A. 90:4324-4327(1993).
RN [10]
RP VARIANTS S-332; S-336; K-354 DEL; D-361; T-392; H-498; Q-510 & W-532.
RX MEDLINE; 94235845.
RA LENZNER C., NUERNBERG P., THIELE B.-J., REIS A., BRABEC V.,
RA SAKALOVA A., JACOBASCH G.;
RT "Mutations in the pyruvate kinase L gene in patients with hereditary
RT hemolytic anemia.";
RL BLOOD 83:2817-2822(1994).
RN [11]
RP VARIANTS E-331; A-341; K-393; S-393; D-458; M-460 AND H-498.
RX MEDLINE; 95221622.
RA BARONCIANI L., BEUTLER E.;
RT "Molecular study of pyruvate kinase deficient patients with
RT hereditary nonspherocytic hemolytic anemia.";
RL J. CLIN. INVEST. 95:1702-1709(1995).
RN [12]
RP VARIANTS HNSHA.
RA BARONCIANI L., WESTWOOD B., BEUTLER E.;
RT "Study of the molecular defects in pyruvate kinase (PK) deficient
RT patients affected by hereditary nonspherocytic hemolytic anemia
RT (HNHA).";
RL J. INVEST. MED. 43:341A-341A(1995).
RN [13]
RP VARIANTS Q-172; Q-337; Q-339; T-357; I-408; T-431; W-486 & Q-532.
RX MEDLINE; 99043610.
RA ZARZA R., ALVAREZ R., PUJADES A., NOMDEDEU B., CARRERA A., ESTELLA J.,
RA REMACHA A., SANCHEZ J.M., MOREY M., CORTES T., PEREZ LUNGMUS G.,
RA BUREO E., VIVES CORRONS J.L.;
RT "Molecular characterization of the PK-LR gene in pyruvate kinase
RT deficient Spanish patients.";
RL BR. J. HAEMATOL. 103:377-382(1998).
RN [14]
RP VARIANTS HNSHA S-332; P-337; W-486; C-498 AND Q-510.
RX MEDLINE; 98141680.
RA PASTORE L., DELLA MORTE R., FRISSO G., ALFINITO F., VITALE D.,
RA CALISE R.M., FERRARO F., ZAGARI A., ROTOLI B., SALVATORE F.;
RT "Novel mutations and structural implications in R-type pyruvate
RT kinase-deficient patients from Southern Italy.";
RL HUM. MUTAT. 11:127-134(1998).
CC -!- CATALYTIC ACTIVITY: ATP + PYRUVATE = ADP + PHOSPHOENOLPYRUVATE.
CC -!- COFACTOR: REQUIRES MAGNESIUM AND POTASSIUM.
CC -!- PATHWAY: FINAL STEP IN GLYCOLYSIS.
CC -!- SUBUNIT: HOMOTETRAMER.
CC -!- DISEASE: DEFECTS IN PYRUVATE KINASE ARE THE MOST COMMON CAUSES
CC FOR HEREDITARY NONSPHEROCYTIC HEMOLYTIC ANEMIA (HNSHA) AMONG
CC GLYCOLYTIC ENZYMES.
CC -!- THERE ARE 4 ISOZYMES OF PYRUVATE KINASE IN MAMMALS: L, R, M1 AND
CC M2. L TYPE IS MAJOR ISOZYME IN THE LIVER, R IS FOUND IN RED
CC CELLS, M1 IS THE MAIN FORM IN MUSCLE, HEART AND BRAIN, AND M2 IS
CC FOUND IN EARLY FETAL TISSUES.
CC -!- ALTERNATIVE PRODUCTS: THE L- AND R-TYPE ISOZYMES ARE PRODUCED
CC FROM A SINGLE GENE BY USE OF DIFFERENT PROMOTERS. THE SEQUENCE
CC SHOWN HERE IS THAT OF THE R-TYPE.
CC -!- SIMILARITY: BELONGS TO THE PYRUVATE KINASE FAMILY.
CC ----------

CC ----------
DR EMBL; D10326; G220036; -.
DR EMBL; M15465; G189996; -.
DR EMBL; S60712; G299641; -.
DR PIR; A30150; A30150.
DR PIR; A29414; A29414.
DR SWISS-2DPAGE; P30613; HUMAN.
DR MIM; 266200; -.
DR PROSITE; PS00110; PYRUVATE_KINASE; 1.
DR PFAM; PF00224; PK; 1.
DR HSSP; P11974; 1PKN.
KW TRANSFERASE; KINASE; GLYCOLYSIS; MULTIGENE FAMILY; PHOSPHORYLATION;
KW MAGNESIUM; ALTERNATIVE SPLICING; DISEASE MUTATION.

```
FT   MOD_RES      43     43       PHOSPHORYLATION (BY CAPK).
FT   ACT_SITE     313    313      BY SIMILARITY.
FT   METAL        315    315      MAGNESIUM (POTENTIAL).
FT   METAL        336    336      MAGNESIUM (POTENTIAL).
FT   METAL        337    337      MAGNESIUM (POTENTIAL).
FT   VARSPLIC      1     33       MSIQENISSLQLRSWVSKSQRDLAKSILIGAPG -> ME
FT                                (IN L-TYPE).
FT   VARIANT      107    107      M -> T (IN HEMOLYTIC ANEMIA).
FT   VARIANT      131    131      MISSING (IN HEMOLYTIC ANEMIA).
FT   VARIANT      134    134      V -> D (IN HEMOLYTIC ANEMIA).
FT   VARIANT      155    155      L -> P (IN HEMOLYTIC ANEMIA).
FT   VARIANT      172    172      E -> Q (IN HEMOLYTIC ANEMIA).
FT   VARIANT      163    163      R -> C (IN HEMOLYTIC ANEMIA; LINZ).
FT   VARIANT      221    221      D -> DD (IN HEMOLYTIC ANEMIA).
FT   VARIANT      275    275      G -> R (IN HEMOLYTIC ANEMIA).
FT   VARIANT      281    281      D -> N (IN HEMOLYTIC ANEMIA).
FT   VARIANT      287    287      F -> V (IN HEMOLYTIC ANEMIA).
FT   VARIANT      314    314      I -> T (IN HEMOLYTIC ANEMIA).
FT   VARIANT      331    331      D -> E (IN HEMOLYTIC ANEMIA).
FT   VARIANT      332    332      G -> S (IN HEMOLYTIC ANEMIA).
FT   VARIANT      336    336      A -> S (IN HEMOLYTIC ANEMIA).
FT   VARIANT      337    337      R -> P (IN HEMOLYTIC ANEMIA).
FT   VARIANT      337    337      R -> Q (IN HEMOLYTIC ANEMIA).
FT   VARIANT      339    339      D -> Q (IN HEMOLYTIC ANEMIA).
FT   VARIANT      341    341      G -> A (IN HEMOLYTIC ANEMIA).
FT   VARIANT      354    354      MISSING (IN HEMOLYTIC ANEMIA).
FT   VARIANT      357    357      I -> T (IN HEMOLYTIC ANEMIA).
FT   VARIANT      359    359      R -> C (IN HEMOLYTIC ANEMIA).
FT   VARIANT      359    359      R -> H (IN HEMOLYTIC ANEMIA).
FT   VARIANT      361    361      N -> D (IN HEMOLYTIC ANEMIA).
FT   VARIANT      368    368      V -> F (IN HEMOLYTIC ANEMIA; OSAKA).
FT   VARIANT      384    384      T -> M (IN HEMOLYTIC ANEMIA; TOKYO/
FT                                BEIRUT; MOST COMMON MUTATION IN JAPANESE
FT                                POPULATION).
FT   VARIANT      392    392      A -> T (IN HEMOLYTIC ANEMIA).
FT   VARIANT      393    393      N -> K (IN HEMOLYTIC ANEMIA).
FT   VARIANT      393    393      N -> S (IN HEMOLYTIC ANEMIA).
FT   VARIANT      401    401      C -> CS (IN HEMOLYTIC ANEMIA).
FT   VARIANT      408    408      T -> I (IN HEMOLYTIC ANEMIA).
FT   VARIANT      421    421      Q -> K (IN HEMOLYTIC ANEMIA; FUKUSHIMA/
FT                                MAEBASHI).
FT   VARIANT      426    426      R -> Q (IN HEMOLYTIC ANEMIA; SAPPORO).
FT   VARIANT      426    426      R -> W (IN HEMOLYTIC ANEMIA).
FT   VARIANT      431    431      A -> T (IN HEMOLYTIC ANEMIA).
FT   VARIANT      458    458      G -> D (IN HEMOLYTIC ANEMIA).
FT   VARIANT      459    459      A -> V (IN HEMOLYTIC ANEMIA).

  Query Match             9.1%;  Score 83;  DB 1;  Length 574;
  Best Local Similarity   26.7%;  Pred. No. 2.2;
  Matches   31;  Conservative   20;  Mismatches   55;  Indels   10;  Gaps    3;

Qy       1 MEEYAREPCPWRIV----DDCGGAFTMGVIGGGVFQAIKGFRNAPVGI----RHRLRGSA 52
           :| :|  |  :| |     |  |     |:: ||    ::  : : | :      | ||:|
Db     138 VESFAGSPLSYRPVAIALDTKGPEIRTGILQGGPESEVELVKGSQVLVTVDPAFRTRGNA 197

Qy      53 NAVRIRAPQIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSG 108
           | | :  | |     | | ::  || ||: |  ::     : :    | || :| |
Db     198 NTVWVDYPNIVRVVPVGGRIY--IDDGLISLVVQKISPEGLVTQVENGGVLGSRKG 251


RESULT  11
PTN5_MOUSE
ID   PTN5_MOUSE     STANDARD;      PRT;   541 AA.
AC   P54830; Q64694;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PROTEIN-TYROSINE PHOSPHATASE STRIATUM-ENRICHED (EC 3.1.3.48) (STEP)
DE   (NEURAL-SPECIFIC PROTEIN-TYROSINE PHOSPHATASE).
GN   PTPN5.
OS   MUS MUSCULUS (MOUSE).
OC   EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC   RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN   [1]
RP   SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC   STRAIN=BALB/C; TISSUE=BRAIN;
RX   MEDLINE; 96115113.
RA   SHARMA E., ZHAO F., BULT A., LOMBROSO P.J.;
RT   "Identification of two alternatively spliced transcripts of STEP: a
RT   subfamily of brain-enriched protein tyrosine phosphatases.";
RL   BRAIN RES. MOL. BRAIN RES. 32:87-93(1995).
CC   -!- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC       PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC   -!- TISSUE SPECIFICITY: STEP20 IS EXPRESSED ONLY IN THE CNS.
CC   -!- ALTERNATIVE PRODUCTS: AT LEAST FOUR FORMS ARE PRODUCED BY
CC       ALTERNATIVE SPLICING: STEP61 (SHOWN HERE), STEP46, STEP38 AND
CC       STEP20. STEP38 AND STEP20 LACK THE CATALYTIC DOMAIN.
CC   -!- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-
CC       TYROSINE PHOSPHATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U28217; G915382; -.
DR   EMBL; U28216; G915380; -.
DR   EMBL; S80329; G1245880; -.
DR   MGD; MGI:97807; PTPN5.
DR   PROSITE; PS00383; TYR_PHOSPHATASE_1; 1.
DR   PROSITE; PS50056; TYR_PHOSPHATASE_2; 1.
DR   PROSITE; PS50055; TYR_PHOSPHATASE_PTP; 1.
DR   PFAM; PF00102; Y_phosphatase; 1.
DR   HSSP; P18052; 1YFO.
KW   HYDROLASE; ALTERNATIVE SPLICING.
FT   DOMAIN       290    533      PROTEIN-TYROSINE PHOSPHATASE.
FT   ACT_SITE     472    472      BY SIMILARITY.
FT   VARSPLIC       1    173      MISSING (IN STEP20 AND STEP46).
FT   VARSPLIC     337    346      GYSGEEKVYI -> VCSSIPRAFH (IN STEP38 AND
FT                                STEP20).
FT   VARSPLIC     347    541      MISSING (IN STEP38 AND STEP20).
SQ   SEQUENCE   541 AA;  60829 MW;  9C9F4460 CRC32;


  Query Match             9.1%;  Score 83;  DB 1;  Length 541;
  Best Local Similarity   28.9%;  Pred. No. 2.1;
  Matches   24;  Conservative   15;  Mismatches   40;  Indels    4;  Gaps    3;

Qy      78 CGLVRLRGKEDPWNSITSGALTGAVLAARS-GPLAMVGSAMMGGILLALIE-GVGILLTR 135
           ||:: | |    |   |:  :: ::      ||:| :||   |   | |:   | :::
Db      76 CGVLWLSGHGHSWLQNTTDLISSSLTVLNHLGPVAWLGSGTWGIPSLLLVSLTVSLVIVT 135

Qy     136 YTAQQFRNAPPFLEDPSQLPPKD 158
                ||| | |: |||:|
Db     136 TLVWHLLKAPP--EPPAPLPPED 156


RESULT  12
GLYA_BORBU
ID   GLYA_BORBU     STANDARD;      PRT;   417 AA.
AC   O51547;
DT   15-DEC-1998 (REL. 37, CREATED)
DT   15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)
DT   15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE   SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE   (SHMT).
GN   GLYA OR BB0601.
OS   BORRELIA BURGDORFERI (LYME DISEASE SPIROCHETE).
OC   BACTERIA; SPIROCHAETALES; SPIROCHAETACEAE; BORRELIA.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE; 98065943.
```

```
RA   FRASER C.M., CASJENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA   LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA   DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA   PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
RA   VAN VUGT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA   UTTERBACK T., WATTHEY L., MCDONALD L., ARTIACH P., BOWMAN C.,
RA   GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA   SMITH H.O., VENTER J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   NATURE 390:580-586(1997).
CC   -!- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC   -!- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC       H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE.
CC   -!- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC       HORMONES AND OTHER COMPONENTS.
CC   -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC   -!- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE001161; G2688518; -.
DR   TIGR; BB0601; -.
DR   PROSITE; PS00096; SHMT; 1.
KW   TRANSFERASE; PYRIDOXAL PHOSPHATE; ONE-CARBON METABOLISM.
FT   BINDING     221    221       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   417 AA;  45881 MW;  D070203A CRC32;

  Query Match             8.8%;  Score 80.5;  DB 1;  Length 417;
  Best Local Similarity   26.4%;  Pred. No. 2.6;
  Matches   38;  Conservative   16;  Mismatches   61;  Indels   29;  Gaps    8;

Qy      13 IVDD------CGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQIGGSF 66
           | ||      |  |   |:|       : || |: : : | |  |      :| |:  |
Db     181 IADDVSAYLLCDIAHIAGLI-------VAGFHNSSIDVAH-LTTSTTHKTLRGPR--GGI 230

Qy      67 AVWGGLFXTIDCGLVRLRGKEDP-WNSITSGALTGAVLAARSGPLAMVGSAMMGGILLAL 125
           : |  |       ||   ||| | :|:: |    |     : |||  |    : |  :|
Db     231 ILSGKDFDK----LVNFNGKEKPLFNAVNSTVFPG----TQGGPLVHV----IAGKAIAF 278

Qy     126 IEGVGILLTRYTAQQFRNAPPFLE 149
           | :     | |   :|     |
Db     279 KEALQESFKEYIANVIKNTKVMAE 302


RESULT  13
Y058_MYCTU
ID   Y058_MYCTU     STANDARD;      PRT;   256 AA.
AC   Q10887;
DT   01-OCT-1996 (REL. 34, CREATED)
DT   01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT   01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE   HYPOTHETICAL 27.8 KD PROTEIN CY251.08.
GN   MTCY251.08.
OS   MYCOBACTERIUM TUBERCULOSIS.
OC   BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC   ACTINOMYCETALES; CORYNEBACTERINEAE; MYCOBACTERIACEAE; MYCOBACTERIUM.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37RV;
RA   CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;
RL   SUBMITTED (JUN-1996) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; Z74410; E249404; -.
KW   HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
FT   TRANSMEM    155    175       POTENTIAL.
FT   TRANSMEM    203    223       POTENTIAL.
SQ   SEQUENCE   256 AA;  27837 MW;  5EDAB7C2 CRC32;

  Query Match             8.8%;  Score 80.5;  DB 1;  Length 256;
  Best Local Similarity   26.2%;  Pred. No. 1.6;
  Matches   39;  Conservative   20;  Mismatches   61;  Indels   29;  Gaps    6;

Qy      11 WRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSA--NAVRIRAPQIG----- 64
           || :  ||  | :|   |        | ||      :|| |   |: |||  :
Db      46 WRCL-RCGD-FALGGPQG---------RGAPEDAPLIMRGKALRQAIIIRALGVERLVRA 94

Qy      64 -----GSFAVW------GGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAM 112
                ::|||       | :  |:|  |  ||      : :|         |||:   ||:
Db      95 LVLALAAWAVWEFRGARGAIQATLDRDLPVLRAAGFKVDQMTVIHALEKALAAKPSTLAL 154

Qy     113 VGSAMMGGILLALIEGVGILLTRYTAQQF 141
           :   :    :|  :||||: | :   : |
Db     155 ITGMLAAYAVLQAVEGVGLWLLKRWGEYF 183


RESULT  14
DC11_DROME
ID   DC11_DROME     STANDARD;      PRT;  1123 AA.
AC   P18169;
DT   01-NOV-1990 (REL. 16, CREATED)
DT   01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT   01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE   DEFECTIVE CHORION-1 FC125 PROTEIN PRECURSOR.
GN   DEC-1.
OS   DROSOPHILA MELANOGASTER (FRUIT FLY).
OC   EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC   PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYDROIDEA;
OC   DROSOPHILIDAE; DROSOPHILA.
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE; 91032553.
RA   WARING G.L., HAWLEY R.J., SCHOENFELD T.;
RT   "Multiple proteins are produced from the dec-1 eggshell gene in
RT   Drosophila by alternative RNA splicing and proteolytic cleavage
RT   events.";
RL   DEV. BIOL. 142:1-12(1990).
CC   -!- FUNCTION: THE DEC-1 FEMALE-STERILE LOCUS PROTEIN IS REQUIRED
CC       FOR PROPER ASSEMBLY OF THE DROSOPHILA EGGSHELL.
CC   -!- ALTERNATIVE PRODUCTS: MULTIPLE PROTEINS ARE PRODUCED FROM THE
CC       DEC-1 GENE BY ALTERNATIVE RNA SPLICING AND PROTEOLYTIC CLEAVAGE
CC       EVENTS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M35887; G157182; -.
DR   PIR; A44766; A44766.
DR   FLYBASE; FBgn0000427; dec-1.
KW   CHORION; EGGSHELL; REPEAT; ALTERNATIVE SPLICING; SIGNAL.
FT   SIGNAL        1     23
FT   CHAIN        24   1123       DEFECTIVE CHORION-1 FC125 PROTEIN.
FT   DOMAIN      493    788       12 X 26 AA TANDEM REPEATS.
```

```
SQ   SEQUENCE   1123 AA;  127959 MW;  99218239 CRC32;


  Query Match             8.8%;  Score 80;  DB 1;  Length 1123;
  Best Local Similarity   24.7%;  Pred. No. 8.3;
  Matches   44;  Conservative   16;  Mismatches   66;  Indels   52;  Gaps   10;

Qy      24 GVIGGGVFQAIKGFRNAPVGI---------RHRLRGSANAVRIRA---------PQIGGS 65
           | :||   |  |   | |  :          ||:  | ::              ||:|
Db     952 GEVGGSESQ--KSNSNPPTTLTPAPQEQPQEHRVHKSPSSAPSETEIENAPSSDPQVGSI 1009

Qy      66 FAVWGGLFXTIDCGLVRLRGKEDPWNS---------ITSGALTGAVLAARSG----PLAM 112
           |   ||      ||: :   :||||           | |    | |  | |       :
Db    1010 FTYGEGLLHPF-MGLLPVERPDDPWNQKPYDPHHPLYTGGGSYDAYL--RDGRHRRDTHI 1066

Qy     113 VGSAMMGGILLALIEGVGIL--LTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSY 168
           :|     |||       |:|  | |      |  |||::  :: |         | ||
Db    1067 MGQGTQHGILTP-----GMLERLLRIKMDFQRRFPPFIQGYAESP---------YESY 1110


RESULT  15
PTGA_BRELA
ID   PTGA_BRELA     STANDARD;      PRT;   674 AA.
AC   Q45298;
DT   01-NOV-1997 (REL. 35, CREATED)
DT   01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT   01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE   PTS SYSTEM, GLUCOSE-SPECIFIC IIABC COMPONENT (EIIABC-GLC) (GLUCOSE-
DE   PERMEASE IIABC COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, ABC
DE   COMPONENT) (EC 2.7.1.69) (EII-GLC / EIII-GLC).
GN   PTSG.
OS   BREVIBACTERIUM LACTOFERMENTUM.
OC   BACTERIA; FIRMICUTES; ACTINOBACTERIA; ACTINOBACTERIDAE;
OC   ACTINOMYCETALES; CORYNEBACTERINEAE; CORYNEBACTERIACEAE;
OC   CORYNEBACTERIUM.
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 13869;
RA   YOON K.H.;
RL   SUBMITTED (NOV-1993) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC   -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOENOLPYRUVATE-DEPENDENT
CC       SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC       -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
CC       AND THE TRANSMEMBRANE CHANNEL; THE IIA DOMAIN CONTAINS THE PRIMARY
CC       PHOSPHORYLATION SITE (THE DONOR IS PHOSPHO-HPR); IIA TRANSFERS ITS
CC       PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC       THE SUGAR.
CC   -!- CATALYTIC ACTIVITY: PROTEIN N-PHOSPHOHISTIDINE + SUGAR =
CC       PROTEIN HISTIDINE + SUGAR PHOSPHATE.
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC   -!- SIMILARITY: CONTAINS A PTS EIIA DOMAIN.
CC   -!- SIMILARITY: CONTAINS A PTS EIIB DOMAIN.
CC   -!- SIMILARITY: CONTAINS A PTS EIIC DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L18875; G405143; -.
DR   PROSITE; PS00371; PTS_EIIA_1; 1.
DR   PROSITE; PS01035; PTS_EIIB_CYS; 1.
DR   PFAM; PF00358; PTS_EIIA_1; 1.
DR   PFAM; PF00367; PTS_EIIB; 1.
DR   HSSP; P08837; 2F3G.
KW   PHOSPHOTRANSFERASE SYSTEM; SUGAR TRANSPORT; TRANSFERASE;
KW   PHOSPHORYLATION; TRANSMEMBRANE.
FT   DOMAIN        1      ?       EIIB DOMAIN.
FT   DOMAIN        ?      ?       EIIC DOMAIN.
FT   DOMAIN        ?    674       EIIA DOMAIN.
FT   TRANSMEM    126    146       POTENTIAL.
FT   TRANSMEM    162    182       POTENTIAL.
FT   TRANSMEM    193    213       POTENTIAL.
FT   TRANSMEM    225    245       POTENTIAL.
FT   TRANSMEM    260    280       POTENTIAL.
FT   TRANSMEM    303    323       POTENTIAL.
FT   TRANSMEM    344    364       POTENTIAL.
FT   TRANSMEM    377    397       POTENTIAL.
FT   TRANSMEM    409    429       POTENTIAL.
FT   TRANSMEM    442    462       POTENTIAL.
FT   TRANSMEM    493    513       POTENTIAL.
FT   MOD_RES      28     28       PHOSPHORYLATION (BY SIMILARITY).
FT   MOD_RES     594    594       PHOSPHORYLATION (BY SIMILARITY).
SQ   SEQUENCE   674 AA;  71626 MW;  C21F0F14 CRC32;


  Query Match             8.7%;  Score 79;  DB 1;  Length 674;
  Best Local Similarity   23.2%;  Pred. No. 6;
  Matches   35;  Conservative   19;  Mismatches   51;  Indels   46;  Gaps    7;

Qy      19 GAFTMGVIGGGVFQAIKGFRN----------------APVGIRHRLRGSANAVRIR-AP 60
           | | :|| | |:   ::   |                 |:|:   |    ||: |:
Db     280 GPFGIGV-GNGISSLLEAVNNFSPFILSIVIPLLYPFLVPLGLHWPL----NAIMIQNLN 334

Qy      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
           :| |              ::|    ||    | :||  | |       :    :||
Db     335 TLGYDF---------------IQGPMGAWNFACFGLVTGVFLIALKEKNRAMRQVSLGG 378

Qy     121 ILLALIEGV------GILLTRYTAQQFRNAP 145
           :|  |: |:      |:|| |:    ||  |
Db     379 MLAGLLGGISEPSLYGVLL-RFKKTYFRLLP 408


Search completed: August 10, 1999, 08:35:42
Job time: 8499 sec
```

RESULT 4
US-08-804-227C-11
; Sequence 11, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rosteck, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: ASCI(DOS) Text only
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,227C
; FILING DATE: February 21, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Plant, Thomas, G.
; REGISTRATION NUMBER: 35,784
; REFERENCE/DOCKET NUMBER: X-8231
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-2459
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1580 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-804-227C-11

Query Match 8.6%; Score 78; DB 2; Length 1580;
Best Local Similarity 27.7%; Pred. No. 2.2;
Matches 39; Conservative 18; Mismatches 50; Indels 34; Gaps 7;

```
Qy      11 WRIVDDC---GGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQIGGSFA 67
           || |      ||   : | |||||    | |   ||:    || :   | |     |
Db    1107 WRRVVGVVAGGGEDQVAVRGGGVF----GRRLVGVGV----RGGSGVWRAR-----GCVV 1153

Qy      68 VWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVG------------ 115
           | |||   :   : |   :    : : :|   | |: |       :||
Db    1154 VTGGL-GGVGGHVARWLARSGAEHVVLAGRRGGGVVGAVELERELVGLGAKVTFVSCDVG 1212

Qy     115 --SAMMGGILLALIEGVGILL 133
             ::|:|  || ::||:|: |
Db    1213 DRASMVG--LLGVVEGLGVPL 1231
```

GenCore version 4.5

OM protein - protein search, using sw model

Run on: August 10, 1999, 06:10:21 ; Search time 9.9 Seconds
(without alignments)
171.451 Million cell updates/sec

Title: US-09-208-619-1
Perfect score: 911
Sequence: 1 MEEYAREPCPWRIVDDCGGA.....QLPPKDGTPAPGYPSYQQYH 172

Scoring table: BLOSUM62

Searched: 106577 seqs, 9868381 residues

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/PCTUS9_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	909	99.8	172	2	US-08-812-645-1	Sequence 1, Appli
2	703.5	77.2	171	2	US-08-812-645-3	Sequence 3, Appli
3	377.5	41.4	158	2	US-08-812-645-4	Sequence 4, Appli
4	78	8.6	1580	2	US-08-804-227C-11	Sequence 11, Appl
5	74	8.1	1891	2	US-08-804-227C-12	Sequence 12, Appl
6	72.5	8.0	977	2	US-08-673-789-8	Sequence 8, Appli
7	72.5	8.0	792	2	US-08-678-039A-40	Sequence 40, Appl
8	72.5	8.0	976	3	PCT-US95-04681-18	Sequence 18, Appl
9	72	7.9	3729	2	US-08-804-227C-4	Sequence 4, Appli
10	71.5	7.8	484	1	US-08-216-276A-17	Sequence 17, Appl
11	68	7.5	2873	2	US-08-466-033-15	Sequence 15, Appl
12	68	7.5	2873	2	US-08-638-911A-2	Sequence 2, Appli
13	68	7.5	853	2	US-08-638-911A-25	Sequence 25, Appl
[redacted]	68	7.5	853	2	US-08-638-911A-27	Sequence 27, Appl
[redacted]	68	7.5	853	2	US-08-638-911A-29	Sequence 29, Appl
16	68	7.5	853	2	US-08-638-911A-31	Sequence 31, Appl
17	68	7.5	357	2	US-08-638-911A-37	Sequence 37, Appl
18	68	7.5	2873	2	US-08-444-733-15	Sequence 15, Appl
19	68	7.5	2873	2	US-08-464-134-15	Sequence 15, Appl
20	68	7.5	2873	2	US-08-461-361-15	Sequence 15, Appl
21	68	7.5	2873	2	US-08-485-910-15	Sequence 15, Appl
22	68	7.5	2873	3	PCT-US95-06266-15	Sequence 15, Appl
23	67	7.4	2910	2	US-08-466-033-183	Sequence 183, App
24	67	7.4	2910	2	US-08-444-733-183	Sequence 183, App
25	67	7.4	2910	2	US-08-464-134-183	Sequence 183, App
26	67	7.4	2910	2	US-08-461-361-183	Sequence 183, App
27	67	7.4	2910	2	US-08-485-910-183	Sequence 183, App
28	67	7.4	2910	3	PCT-US95-06266-157	Sequence 157, App
29	66	7.2	1163	1	US-08-173-497-4	Sequence 4, Appli
30	66	7.2	1163	1	US-08-286-889-4	Sequence 4, Appli
31	66	7.2	1163	1	US-08-485-618-4	Sequence 4, Appli
32	66	7.2	1163	2	US-08-362-652-4	Sequence 4, Appli
33	66	7.2	1163	2	US-08-605-672-4	Sequence 4, Appli
34	66	7.2	1163	2	US-08-482-293A-4	Sequence 4, Appli
35	66	7.2	1163	2	US-08-943-363-4	Sequence 4, Appli
36	66	7.2	529	2	US-08-370-909-19	Sequence 19, Appl
37	66	7.2	529	2	US-08-504-048-8	Sequence 8, Appli
38	66	7.2	1163	2	US-08-476-062A-44	Sequence 44, Appl
39	66	7.2	1163	3	PCT-US96-01314-44	Sequence 44, Appl
40	65	7.1	196	2	US-08-467-822-42	Sequence 42, Appl
41	65	7.1	199	2	US-08-467-822-43	Sequence 43, Appl
42	64.5	7.1	738	1	US-07-985-458-3	Sequence 3, Appli
43	64.5	7.1	605	1	US-08-462-884A-3	Sequence 3, Appli
44	64.5	7.1	560	1	US-07-891-942G-5	Sequence 5, Appli
45	64.5	7.1	605	2	US-08-461-881B-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT   1
US-08-812-645-1
; Sequence 1, Application US/08812645
; Patent No. 5854414
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN MITOCHONDRIAL
;    TITLE OF INVENTION:  MEMBRANE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/812,645
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  To Be Assigned
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0229 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  172 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      LIBRARY:  BLADNOT04
;      CLONE:  1318463
US-08-812-645-1


  Query Match             99.8%;  Score 909;  DB 2;  Length 172;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-97;
  Matches  172;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MEEYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAP 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MEEYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAP 60

Qy      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120

Qy     121 ILLALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSYQQYH 172
```

```
                     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ILLALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSYQQYH 172


RESULT   2
US-08-812-645-3
; Sequence 3, Application US/08812645
; Patent No. 5854414
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN MITOCHONDRIAL
;    TITLE OF INVENTION:  MEMBRANE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/812,645
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  To Be Assigned
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0229 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  171 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  GI 1770564
US-08-812-645-3


  Query Match           77.2%;  Score 703.5;  DB 2;  Length 171;
  Best Local Similarity  75.4%;  Pred. No. 6.1e-74;
  Matches  129;  Conservative   17;  Mismatches   24;  Indels    1;  Gaps    1;

Qy       1 MEEYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAP 60
           |||||||||||||||||||||||| ||||:|||||||||:|||: ||||||  |:: |||
Db       1 MEEYAREPCPWRIVDDCGGAFTMGTIGGGIFQAIKGFRNSPVGVNHRLRGSLTAIKTRAP 60

Qy      61 QIGGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGG 120
           |:||||||||||  ||| :|::|||||||||||||||||:||||:||:|||||| |||
Db      61 QLGGGSFAVWGGLFSMIDCSMVQVRGKEDPWNSITSGALTGAILAARNGPVAMVGSAAMGG 120

Qy     121 LLALIEGVGILLTRYTAQQFRNAPPFLEDPSQLPPKDGTPAPGYPSYQQY 171
           |||||||| ||||||: : || | | | |||||||     |:  :  |:||
Db     121 LLALIEGAGILLTRFASAQFPNGPQFAEDPSQLPSTQ-LPSSPFGDYRQY 170


RESULT   3
US-08-812-645-4
; Sequence 4, Application US/08812645
; Patent No. 5854414
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Goli, Surya K.
;    TITLE OF INVENTION:  NOVEL HUMAN MITOCHONDRIAL
;    TITLE OF INVENTION:  MEMBRANE PROTEIN
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/812,645
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  To Be Assigned
;      FILING DATE:  Filed Herewith
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0229 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  158 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  GI 557267
US-08-812-645-4


  Query Match           41.4%;  Score 377.5;  DB 2;  Length 158;
  Best Local Similarity  48.7%;  Pred. No. 2.4e-36;
  Matches   75;  Conservative   26;  Mismatches   50;  Indels    3;  Gaps    2;

Qy       3 EYAREPCPWRIVDDCGGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQI 62
           :::|:|||  |::| |||| || ||| |:  ||||||:|:|    |  |: :|:: ||| :
Db       4 DHSRDPCPIVILNDFGGAFAMGAIGGVVWHGIKGFRNSPLG--ERGSGAMSAIKARAPVL 61

Qy      63 GGSFAVWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVGSAMMGGIL 122
           ||:| |||||| | || :  :| :|||||:| :|  ||  || | |      |::    |
Db      62 GGNFGVWGGLFSTFDCAVKAVRKREDPWNAIIAGFFTGGALAVRGGWRHTRNSSITCACL 121

Qy     123 LALIEGVGILLTRYTAQQFR-NAPPFLEDPSQLP 155
           | :||||::  || | | :  |||  | ||  |
Db     122 LGVIEGVGLMFQRYAAWQAKPMAPPLPEAPSSQP 155


RESULT   4
US-08-804-227C-11
; Sequence 11, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
```

```
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1580 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-227C-11


  Query Match             8.6%;  Score 78;  DB 2;  Length 1580;
  Best Local Similarity   27.7%;  Pred. No. 2.2;
  Matches   39;  Conservative   18;  Mismatches   50;  Indels   34;  Gaps    7;

Qy      11 WRIVDDC---GGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQIGGSFA 67
           || |       ||   : | |||||    | |   ||:    || :   | |     |
Db    1107 WRRVVGVVAGGGEDQVAVRGGGVF----GRRLVGVGV----RGGSGVWRAR-----GCVV 1153

Qy      68 VWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVG---------- 115
           | |||    :   : |   :    : : :|   | |: |      :||
Db    1154 VTGGL-GGVGGHVARWLARSGAEHVVLAGRRGGGVVGAVELERELVGLGAKVTFVSCDVG 1212

Qy     115 --SAMMGGILLALIEGVGILL 133
             ::|:|  || ::||:|: |
Db    1213 DRASMVG--LLGVVEGLGVPL 1231


RESULT   5
US-08-804-227C-12
; Sequence 12, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1891 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  peptide
US-08-804-227C-12


  Query Match             8.1%;  Score 74;  DB 2;  Length 1891;
  Best Local Similarity   27.0%;  Pred. No. 8.3;
  Matches   38;  Conservative   19;  Mismatches   50;  Indels   34;  Gaps    7;

Qy      11 WRIVDDC---GGAFTMGVIGGGVFQAIKGFRNAPVGIRHRLRGSANAVRIRAPQIGGSFA 67
           || |       ||   : | |||||    | |   ||:    || :   | |     |
Db    1192 WRRVVGVVAGGGEDQVAVRGGGVF----GRRLVGVGV----RGGSGVWRAR-----GCVV 1238

Qy      68 VWGGLFXTIDCGLVRLRGKEDPWNSITSGALTGAVLAARSGPLAMVG---------- 115
           | |||    :   : |   :    : : :|   | |: |      :||
Db    1239 VTGGL-GGVGGHVARWLARSGAEHVVLAGRRGGGVVGAVELERELVGLGAKVTFVSCDVG 1297

Qy     115 --SAMMGGILLALIEGVGILL 133
             ::::|  || ::||:|: |
Db    1298 DRASVVG--LLGVVEGLGVPL 1316


RESULT   6
US-08-673-789-8
; Sequence 8, Application US/08673789
; Patent No. 5814479
;  GENERAL INFORMATION:
;    APPLICANT:  ZHOU, RENPING; SCHULZ, NICHOLAS,
;    APPLICANT:  T.; KROMER, LAWRENCE, F.; VANDE WOUDE,
;    APPLICANT:  GEORGE, F.
;    TITLE OF INVENTION:  BSK RECEPTOR LIKE
;    TITLE OF INVENTION:  TYROSINE KINASE AND LIGAND AND THEIR
;    TITLE OF INVENTION:  USE IN DIAGNOSTIC AND THERAPEUTIC
;    TITLE OF INVENTION:  METHODS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/673,789
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/177,812
;      FILING DATE:  04-JAN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CAROL M. GRUPPI
;      REGISTRATION NUMBER:  37,341
;      REFERENCE/DOCKET NUMBER:  2026-4105
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
```

; TELEFAX: (212) 751-6849
; TELEX: 421792
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 977
; TYPE: AMINO ACID
; STRANDEDNESS: UNKNOWN
; TOPOLOGY: UNKNOWN
US-08-673-789-8

Query Match 8.0%; Score 72.5; DB 2; Length 977;
Best Local Similarity 23.6%; Pred. No. 4.9;
Matches 29; Conservative 22; Mismatches 53; Indels 19; Gaps 5;

Qy 44 IRHRLRGSANAVRIRAPQIGGSFAVWGGLFXTID------CGLVRLRGKEDPWNSITSGA 97
Db 472 VTYRKKGDSNSYNVRRTE---------GFSVTLDDLAPDTTYLVQVQALTQEGQGAGSKV 522

Qy 98 LTGAVLAAR-SGPLAMVGSAMMGGILLALIEGVGILL-TRYTAQQFRNAPP--FLEDPSQ 153
D[redacted] 523 HEFQTLSPEGSGNLAVIGGVAVGVVLLLVLAGVGFFIHRRRKNQRARQSPEDVYFSKSEQ 582

Qy 154 LPP 156
Db 583 LKP 585

RESULT 7
US-08-678-039A-40
; Sequence 40, Application US/08678039A
; Patent No. 5858662
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Morris, Colleen A.
; TITLE OF INVENTION: Diagnosis of Williams Syndrome and
; TITLE OF INVENTION: Williams Syndrome Cognitive Profile by Analysis of the
; TITLE OF INVENTION: Presence or Absence of a LIM-Kinase Gene
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; STREET: Tower
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; [redacted] MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/678,039A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-120A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-624-1589
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 792 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-678-039A-40

Query Match 8.0%; Score 72.5; DB 2; Length 792;
Best Local Similarity 27.0%; Pred. No. 3.7;
Matches 34; Conservative 19; Mismatches 48; Indels 25; Gaps 6;

Qy 24 GVIGGGVFQAIKGF-RNAPVGIRHRL-----RGSANAVR-----IRAPQIGGSFAV---- 69
Db 602 GVPGLGVGAGVPGFGAGADEGVRRSLSPELREGDPSSSQHLPSTPSSPRVPGALAAAKAA 661

Qy 69 -WGGLFXTIDCGLVRLRGKEDPWNSITSG---ALTGAVLAARSGPLAMVGSAMMGGILLA 124
Db 662 KYGAAVPGVLGGLGALGGVGIPGGVVGAGPAAAAAAAKAAAKAAQFGLVGAAGLGGL--- 719

Qy 125 LIEGVG 130
Db 719 ---GVG 721

RESULT 8
PCT-US95-04681-18
; Sequence 18, Application PC/TUS9504681
; GENERAL INFORMATION:
; APPLICANT: Fox, Gary M.
; TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amgen Patent Operations/RBW
; STREET: 1840 Dehavilland Drive
; CITY: Thousand Oaks
; STATE: California
; COUNTRY: USA
; ZIP: 91320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04681
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Winter, Robert B.
; REFERENCE/DOCKET NUMBER: A-287
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 976 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04681-18

Query Match 8.0%; Score 72.5; DB 3; Length 976;
Best Local Similarity 23.6%; Pred. No. 4.9;
Matches 29; Conservative 22; Mismatches 53; Indels 19; Gaps 5;

Qy 44 IRHRLRGSANAVRIRAPQIGGSFAVWGGLFXTID------CGLVRLRGKEDPWNSITSGA 97
Db 471 VTYRKKGDSNSYNVRRTE---------GFSVTLDDLAPDTTYLVQVQALTQEGQGAGSKV 521

Qy 98 LTGAVLAAR-SGPLAMVGSAMMGGILLALIEGVGILL-TRYTAQQFRNAPP--FLEDPSQ 153
Db 522 HEFQTLSPEGSGNLAVIGGVAVGVVLLLVLAGVGFFIHRRRKNQRARQSPEDVYFSKSEQ 581

Qy 154 LPP 156
Db 582 LKP 584

RESULT 9
US-08-804-227C-4
; Sequence 4, Application US/08804227C

```
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3729 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-08-804-227C-4


  Query Match             7.9%;  Score 72;  DB 2;  Length 3729;
  Best Local Similarity   25.6%;  Pred. No. 36;
  Matches   45;  Conservative   16;  Mismatches   63;  Indels   52;  Gaps    6;

Qy      47 RLRGSANAVRIRAP--------QIGG----SFAVWGGLFXTIDCGLVRLRGKEDPWNSIT 94
           |||||   | :           | ||     :   |     :   |  : | | |  ::
Db    1666 RLRGSRTGVFVGTNGQHYMPLLQNGGDSFDGYLGTGNSASVMSGRLSYVFGLEGPAVTVD 1725

Qy      95 SG------ALTGAVLAARSG--PLAMVGSAMM-------------------------G 119
           :       ||  || | | |   :|:|| | :                         |
Db    1726 TACSASLVALHLAVQAMRRGECDMALVGGATVMSTPEMLVEFSRQRVISANGRSRAFAAG 1785

Qy     120 GILLALIEGVGILLTRYTAQQFRNAPPFL-----EDPSQLPPKDGTPAPGYPSYQQ 170
           :|| |||||:||    :   ||  | |        :|    :|  ||  || |:
Db    1786 ADGVALGEGVGVLLVERLSDAERNGHPVLAVVRGSAVNQDGASNGLTAPNGPSQQR 1841


RESULT  10
US-08-216-276A-17
; Sequence 17, Application US/08216276A
; Patent No. 5595912
;  GENERAL INFORMATION:
;    APPLICANT:  VAKHARIA, VIKRAM
;    APPLICANT:  SNYDER, DAVID
;    TITLE OF INVENTION:  SPECIFIC DNA AND RNA SEQUENCES
;    TITLE OF INVENTION:  ASSOCIATED WITH US IBDV VARIANTS, VECTOR CARRYING DNA
;    TITLE OF INVENTION:  SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID
;    TITLE OF INVENTION:  SEQUENCES, VACCINE AND METHOD OF VACCINATION
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 S. Jefferson Davis Highway, Suite 400
;      CITY:  Arlington
;      STATE:  Virginia
;      COUNTRY:  U.S.A.
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/216,276A
;      FILING DATE:  23-MAR-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/083,784
;      FILING DATE:  28-JUN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/519,202
;      FILING DATE:  04-MAY-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/227,311
;      FILING DATE:  02-AUG-1988
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kelber, Steven B.
;      REGISTRATION NUMBER:  30,073
;      REFERENCE/DOCKET NUMBER:  2747-054-27 CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;      TELEX:  248855 OPAT UR
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  484 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-216-276A-17


  Query Match             7.8%;  Score 71.5;  DB 1;  Length 484;
  Best Local Similarity   22.4%;  Pred. No. 2.4;
  Matches   41;  Conservative   26;  Mismatches   73;  Indels   43;  Gaps    9;

Qy      22 TMGVIGGGVFQAIKGFRNAPVGIRHRLRGSAN----AVRIRAPQIGGSFAVWGGL----- 73
           |:|  | |:     ||  : ||  : |: : |     : : |  :  |:  |  :
Db      59 TVGDTGSGLIVFFPGFPGSIVGAHYTLQSNGNYKFDQMLLTAQNLPASYNYWRLVSWSLT 118

Qy      73 --FXTIDCGLVRLRGKEDPWNSIT-SGALT-------GAVLAARSGPLAMVGSAMMGGIL 122
             |:  ||  | |     |::|  |:|:          :::| :     :|: ::|
Db     119 VRSSTVPGGLYALNG---TINAVTFQGSLSELTDVSYNGLMSATANINDKIGNVLVG--- 173

Qy     123 LALIEGVGILL--TRYTAQQFRNAPPFLE---DPSQLPPKDG---------TPAPGYPSY 168
           |||  :|   | |     |   |      ||  :   |           | |  |
Db     173 ----EGVTVLSLPTSYDVGYVRLGDPIPAIGLDPKMVATCDSSDRPRVYTITAADNYQFS 228

Qy     169 QQY 171
           ||
Db     229 SQY 231


RESULT  11
US-08-466-033-15
; Sequence 15, Application US/08466033
; Patent No. 5766840
;  GENERAL INFORMATION:
;    APPLICANT:  Kim, Jungsuh P.
;    APPLICANT:  Wages, John
;    APPLICANT:  Young, LaVonne M.
;    APPLICANT:  Fry, Kirk E.
;    APPLICANT:  Linnen, Jeffrey M.
;    TITLE OF INVENTION:  Hepatitis G Virus and Molecular
;    TITLE OF INVENTION:  Cloning Thereof
;    NUMBER OF SEQUENCES:  277
;    CORRESPONDENCE ADDRESS:
```

; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,033
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/389,886
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/357,509
; FILING DATE: 16-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/329,729
; FILING DATE: 26-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/344,271
; FILING DATE: 23-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,558
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/285,543
; FILING DATE: 03-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/246,985
; FILING DATE: 20-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-0201.36/G100P11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2873 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
U[redacted]466-033-15

Query Match 7.5%; Score 68; DB 2; Length 2873;
Best Local Similarity 29.5%; Pred. No. 73;
Matches 18; Conservative 13; Mismatches 24; Indels 6; Gaps 2;

Qy 107 SGPLAMVGSAMMGGILLALIEGVGILLTRYT-----AQQFRNAPPFLEDP-SQLPPKDGT 160
:||: |:| |: ||:: | | :::| : | :|: | |:|| |
Db 1553 AGPILMIGLAIAGGMIYASYTGSLVVVTDWDVKGGGAPLYRHGDQATPQPVVQVPPVDHR 1612

Qy 161 P 161
|
Db 1613 P 1613

RESULT 12
US-08-638-911A-2
; Sequence 2, Application US/08638911A
; Patent No. 5766916
; GENERAL INFORMATION:
; APPLICANT: Belaeyv, Alexander S.
; APPLICANT: Chong, Susan M.
; TITLE OF INVENTION: Hepatitis G Virus Protease
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/638,911A
; FILING DATE: 04/24/96
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, Susan T.
; REGISTRATION NUMBER: 38,443
; REFERENCE/DOCKET NUMBER: 4600-0124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2873 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-638-911A-2

Query Match 7.5%; Score 68; DB 2; Length 2873;
Best Local Similarity 29.5%; Pred. No. 73;
Matches 18; Conservative 13; Mismatches 24; Indels 6; Gaps 2;

Qy 107 SGPLAMVGSAMMGGILLALIEGVGILLTRYT-----AQQFRNAPPFLEDP-SQLPPKDGT 160
:||: |:| |: ||:: | | :::| : | :|: | |:|| |
Db 1553 AGPILMIGLAIAGGMIYASYTGSLVVVTDWDVKGGGAPLYRHGDQATPQPVVQVPPVDHR 1612

Qy 161 P 161
|
Db 1613 P 1613

RESULT 13
US-08-638-911A-25
; Sequence 25, Application US/08638911A
; Patent No. 5766916
; GENERAL INFORMATION:
; APPLICANT: Belaeyv, Alexander S.
; APPLICANT: Chong, Susan M.
; TITLE OF INVENTION: Hepatitis G Virus Protease
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/638,911A
; FILING DATE: 04/24/96
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Evans, Susan T.

```
;       REGISTRATION NUMBER:  38,443
;       REFERENCE/DOCKET NUMBER:  4600-0124
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 324-0880
;       TELEFAX:  (415) 324-0960
;   INFORMATION FOR SEQ ID NO:  25:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  853 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-638-911A-25


  Query Match             7.5%;  Score 68;  DB 2;  Length 853;
  Best Local Similarity   29.5%;  Pred. No. 14;
  Matches   18;  Conservative   13;  Mismatches   24;  Indels    6;  Gaps    2;

Qy     107 SGPLAMVGSAMMGGILLALIEGVGILLTRYT-----AQQFRNAPPFLEDP-SQLPPKDGT 160
           :||: |:| |: ||:: |   |  :::| :      |  :|:       |  |:|| |
Db     748 AGPILMIGLAIAGGMIYASYTGSLVVVTDWDVKGGGAPLYRHGDQATPQPVVQVPPVDHR 807

Qy     161 P 161
           |
Db     808 P 808


RESULT  14
US-08-638-911A-27
; Sequence 27, Application US/08638911A
; Patent No. 5766916
;  GENERAL INFORMATION:
;    APPLICANT:  Belaeyv, Alexander S.
;    APPLICANT:  Chong, Susan M.
;    TITLE OF INVENTION:  Hepatitis G Virus Protease
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/638,911A
;      FILING DATE:  04/24/96
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Susan T.
;      REGISTRATION NUMBER:  38,443
;      REFERENCE/DOCKET NUMBER:  4600-0124
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  27:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  853 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-638-911A-27


  Query Match             7.5%;  Score 68;  DB 2;  Length 853;
  Best Local Similarity   29.5%;  Pred. No. 14;
  Matches   18;  Conservative   13;  Mismatches   24;  Indels    6;  Gaps    2;

Qy     107 SGPLAMVGSAMMGGILLALIEGVGILLTRYT-----AQQFRNAPPFLEDP-SQLPPKDGT 160
           :||: |:| |: ||:: |   |  :::| :      |  :|:       |  |:|| |
Db     748 AGPILMIGLAIAGGMIYASYTGSLVVVTDWDVKGGGAPLYRHGDQATPQPVVQVPPVDHR 807

Qy     161 P 161
           |
Db     808 P 808


RESULT  15
US-08-638-911A-29
; Sequence 29, Application US/08638911A
; Patent No. 5766916
;  GENERAL INFORMATION:
;    APPLICANT:  Belaeyv, Alexander S.
;    APPLICANT:  Chong, Susan M.
;    TITLE OF INVENTION:  Hepatitis G Virus Protease
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Dehlinger & Associates
;      STREET:  350 Cambridge Avenue, Suite 250
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94306
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/638,911A
;      FILING DATE:  04/24/96
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Susan T.
;      REGISTRATION NUMBER:  38,443
;      REFERENCE/DOCKET NUMBER:  4600-0124
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  29:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  853 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-638-911A-29


  Query Match             7.5%;  Score 68;  DB 2;  Length 853;
  Best Local Similarity   29.5%;  Pred. No. 14;
  Matches   18;  Conservative   13;  Mismatches   24;  Indels    6;  Gaps    2;

Qy     107 SGPLAMVGSAMMGGILLALIEGVGILLTRYT-----AQQFRNAPPFLEDP-SQLPPKDGT 160
           :||: |:| |: ||:: |   |  :::| :      |  :|:       |  |:|| |
Db     748 AGPILMIGLAIAGGMIYASYTGSLVVVTDWDVKGGGAPLYRHGDQATPQPVVQVPPVDHR 807

Qy     161 P 161
           |
Db     808 P 808


Search completed: August 10, 1999, 08:35:21
Job time: 8700 sec
```